**REMARKS**

The Examiner's suggestion to amend the title has been accepted. The case has also been put proper U.S. format. The drawing objections have been overcome by labeling the boxes of the drawings, as suggested by the Examiner.

With regard to the claims, the independent claims have each been amended to more clearly differentiate the subject matter from the two prior art documents by incorporating the wording of claim 4. In particular the feature that the selection of the key frames which are extracted from the plurality of programs is made in response to the user's input and is performed interactively, has been added. For example, the key frames extracted may be selected to be only those from received programs which have actually been selected to be viewed by the user via the apparatus. In addition, or alternatively, the key frames are selected from those programs which fall within a category or categories of programs which the user has informed the system they are interested in.

In the current application no claim is made to the fact that the extraction of key frames is itself inventive and it is acknowledged in the current application in paragraph 0025 that frame extraction is known. It is the selection of the specific key frames as a result of user interaction and then the subsequent use of the key frames that is important and inventive.

In the two prior art documents used by the Examiner there is no disclosure of the current invention as now defined, either alone or in combination.

In the Ismail document there is disclosed a system for revising a user profile based on certain data, typically that which is generated via an EPG system. However, and as the examiner acknowledges in their report, there is no disclosure of the use of the extraction of key frames from the received video. Furthermore, there is no requirement to do so, or anything to suggest to

the skilled person that there would be such a need as the recommendations can be generated in accordance with the Ismail scheme adequately using the steps set out in that patent. Ismail therefore discloses a different, and independent scheme for generating recommendations.

The Wilf document discloses another independent scheme which primarily is to provide an informed and more accurate response to a user's query for information or program details and for the response which is generated to be focused and of relevance to the query. Wilf does disclose the possibility to capture key frames of video but appears to only disclose that the system captures key frames from "all" received video data. The examiner points in his objection to the claim 4 wording of the current application, to paragraphs 0038, 0044, 0045 of the Wilf document, but, with specific reference to paragraph 0045 and the Figure 3, it appears that key frames are taken for all received video content which is received on the audio video stream. In contrast in the current invention key frames are only taken for video which is selected in accordance with the user's interactive input.

It is therefore submitted that, firstly there is no suggestion in Ismail to the skilled person that they should or could combine the teaching of Wilf therewith as an obvious step and, secondly, even if they did, the skilled person would still not arrive at the invention of the current application, as now defined, as Wilf does not disclose the selection of key frames based on the user's interaction with the system.

**IDS Prior Art**

Enclosed per the Examiner's comments are copies of Smeaton et al., "The Fischlár Digital Video System: A Digital Library of Broadcast TV Programmes"; and O'Connor et al., "Fischlár:

An On-line System for Indexing and Browsing of Broadcast Television Content". Both of these were cited in the preliminary IDS. They are believed to be only background art.

Applicant solicits favorable action in view of the amendments and arguments here presented. Corrected drawings have also been submitted.

**Conclusion**

This is a request under the provision of 37 CFR § 1.136(a) to extend the period for filing a response in the above-identified application for three months from June 19, 2008 to September 19, 2008. Applicant is a large entity; therefore, please charge Deposit Account number 26-0084 in the amount of $1,050.00 to cover the cost of the three-month extension. Any deficiency or overpayment should be charged or credited to Deposit Account 26-0084.

No fees or extensions of time are believed to be due in connection with this amendment; however, consider this a request for any extension inadvertently omitted, and charge any additional fees to Deposit Account No. 26-0084.

Reconsideration and allowance is respectfully requested.

Respectfully submitted,

EDMUND J. SEASE, Reg. No. 24741
McKEE, VOORHEES & SEASE, P.L.C.
801 Grand Avenue, Suite 3200
Des Moines, Iowa 50309-2721
Phone No: (515) 288-3667
Fax No: (515) 288-1338
**CUSTOMER NO: 22885**

Attorneys of Record

- pw -

XP-002271605

P.D. 00.00.2001
P. 312-313
2

# The Físchlár Digital Video System: A Digital Library of Broadcast TV Programmes

A. F. Smeaton, N. Murphy, N. E. O'Connor, S. Marlow,
H. Lee, K. McDonald, P. Browne, and J. Ye
Centre for Digital Video Processing
Dublin City University, Glasnevin, Dublin 9, Ireland
+353 – 1 - 7005262
Alan.Smeaton@dcu.ie

## ABSTRACT

Físchlár is a system for recording, indexing, browsing and playback of broadcast TV programmes which has been operational on our University campus for almost 18 months. In this paper we give a brief overview of how the system operates, how TV programmes are organised for browse/playback and a short report on the system usage by over 900 users in our University.

## 1. INTRODUCTION

The Físchlár digital video system is a web-based system for recording, analysis, browsing and playback of TV programmes which is used within the campus environment at Dublin City University. It allows users to initiate the recording of programmes from any of the 8 terrestrial TV stations for our area. Once digitised, programmes are analysed for shot boundaries and shot-based representative frames are selected [1]. Shots are then clustered into scenes using a variety of techniques. Físchlár is accessed through a conventional web browser on a desktop machine although we have developed a WAP interface to allow users to reserve programme recordings through a mobile phone and we developed a version of the browser/player for a PDA in a mobile environment. The Físchlár system is described in [2].

In its current configuration, Físchlár allows users to record, play (stream) and browse programmes, and it is the way we allow browsing of TV programme content which is one of the reasons that Físchlár is novel. Commercially available TV recording devices such as TiVo [3] allow recording and playback of broadcast TV content but the browsing function is little more than fast forward and rewind. In Físchlár we have developed 8 different browser interfaces, described and evaluated elsewhere [4], each of which is tailored to a user's task, context and preferences. For example, there are different keyframe browsers for users depending upon whether they have seen the particular programme being browsed before and are interested in locating a particular scene they know, or they are watching for the first time.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
*JCDL'01*, June 24-28, 2001, Roanoke, Virginia, USA.
Copyright 2001 ACM 1-58113-345-6/01/0006...$5.00.

There are browsers for users who prefer linear vs. structured browsing or for users who prefer a static vs. a dynamic interaction style. Altogether this means that there is sound support for a user's specific task and the context for their need.

Another aspect of Físchlár which is novel, and which is of interest here, is that it has been integrated with a large-scale TV recommender system called PTV [5] and in the remainder of this paper we give a brief overview of what that provides and how it is being used.

## 2. THE COMBINED FÍSCHLÁR-PTV SYSTEM

Físchlár has been in continuous use for almost 18 months and during that time we have developed, refined and enhanced its functionality. Digital video, in MPEG-1 format, is stored on a SUN Enterprise video server and our archive has over 300 hours of TV (about 400 broadcast programmes) content at any one time. Our server is capable of streaming to over 200 clients concurrently via a web browser plug-in and the system is available from student residences, undergraduate and postgraduate laboratories, and from the main library on campus.

All users must register before using Físchlár and through a logging on process, we are able to track usage and offer personalised services. This includes remembering the user's favoured browser interface as well as programmes. Users use the system mostly for entertainment but also for study-related activities such as browsing/playback of news or specialist programmes (broadcast documentaries, etc.).

Físchlár has two modes of operation, one for recording and another for browse/playback. In recording mode, users are presented with TV listings for the 8 major terrestrial TV channels within our area, for today and for tomorrow. We provide, for each programme, some text details on what the programme is about, who it stars etc., taken from an online entertainment guide. Programmes are also automatically assigned to one or more of a dozen genres including sport, documentary, soap, movies, music, kids, home and garden, etc. Users can view the programme listings by TV station, or across the broadcast channels by genre.

Each Físchlár user is also an indirect user of the PTV system. PTV generates recommendations of TV programmes for users to watch based on their past preferences (positive and negative), the past preferences of others who share or have differing preferences, and the descriptions of new, unviewed programmes. PTV uses

case based reasoning as part of its underlying processing and this is described in [5].

A transparent link between Fischlár and PTV allows each Fischlár user's TV viewing recommendations, from PTV, to be presented alongside the TV listings by channel and by genre. In this way we provide not only the standard and genre-organised TV listings for the next 2 days but also a personalised view of what programmes PTV thinks a user should explicitly request recording of. The function of the recording mode in Fischlár is to have the user explicitly select TV programmes which s/he wants to be recorded and to invite the user to grade each programme on a scale of 1 to 5 in terms of their interest in viewing it. These ratings are then passed back to PTV leading to higher quality recommendations. The alternative to having users explicitly request program recording is recording 24/7 on all 8 TV channels but this would allow us to maintain an archive of only the recently broadcast programmes. We feel that users may prefer to access not just materials from within the last week but further back in time, and selective recording rather than taking a shotgun approach, allows us to do just that.

In browse/playback mode, each user is presented with the full library of recorded programmes among which to browse, as well as our within-programme browsing facilities based on keyframe navigation. Recorded programmes (at any one time over 300 hours) can be viewed by TV station, by genre, or by examining recordings from the most recent 7 days only. In addition, we use the connection with PTV to allow personalised recommendations of programmes from the archive to be viewed as well as a category called "favourites", corresponding to subsequent episodes of a user's previous viewings. From a user's perspective this means that when using Fischlár, a user is presented with a TV schedule for the next 2 days from which s/he can request specific programmes to be recorded, with personalised recommendations built in, and a user can browse an archive or library of already broadcast and recorded programmes, again with personalised recommendations built in.

On selecting a specific programme, a user is immediately presented with the set of keyframes drawn from that program which can be a large number of images. For example, a recent 25 minute episode of "The Simpsons" generated 313 keyframes, a 50 minute episode of "Little House on the Prairie" generated 326 and the movie "Crimson Tide", which is 2 hours and 5 minutes, generated 1755 keyframes. Our different browser interfaces described in [4] are used to provide efficient navigation through these keyframes. As a user is browsing the keyframes, he/she can switch to streaming the video of that programme from that keyframe onwards, by clicking the keyframe.

## 3. USAGE OF THE COMBINED FÍSCHLÁR-PTV SYSTEM

At the time of writing there are over 900 registered users of the Fischlár system. Some of these users are using old PCs in residences, donated by the University to the project and others are using their own desktop machines from within the University intranet. Almost 3000 recording requests have been received in the last 12 months, with 1034 of these requests by 105 users in the last 2 months alone. In fact with our video server limited to storing 300 hours only, we have to remove programmes older than about 1 month in order to provide space for incoming material. The programmes most frequently recorded are "The Simpsons" and "Friends". Other popular programmes are Star Trek (SciFi), Top of the Pops (music), Coronation Street (soap) and 100 years (documentary).

Almost 30% of our users have logged into the system 5 times or more but this statistic is a bit misleading since a single login persists over the whole of a browser's session, so if a user "logs on" to the system then that session remains until the browser application on the PC is shut down. Feedback from users has been hugely positive, especially from those using it in residences. Using the system from labs is less comfortable for users and has been likened to watching TV in public.

In coupling Fischlár with the PTV system we have extended the usefulness of both systems and the combined system presents the user with personalised access to a library of digital video materials. We will shortly introduce other functionality to the system such as text-searching and content-based alerting based on teletext capture. One application which has been requested by staff and students is what we call "buddy clipping", the ability to scope out and define a clip of video from the library whose address can be emailed as an embedded link to others, with text annotation. We have also extended the Fischlár interface to operate on a Compaq iPAQ, a mobile PDA which accesses the system over a wireless LAN.

## 4. REFERENCES


[1] Evaluating and Combining Digital Video Shot Boundary Detection Algorithms. Browne P, Smeaton A, Murphy N, O'Connor N, Marlow S and Berrut C. In Proceedings of IMVIP 2000, Belfast, Northern Ireland, September 2000.

[2] O'Connor, N., Marlow, S., Murphy, N., Smeaton, A., Browne, P., Deasy, S., Lee, H. and Mc Donald, K. Fischlár: an On-line System for Indexing and Browsing of Broadcast Television Content. In Proceedings of *ICASSP 2001* (Salt Lake City, UT, May, 2001).

[3] TiVo. http://www.tivo.com.

[4] User Interface Design for Keyframe-Based Content Browsing of Digital Video. Lee., H. PhD thesis, Dublin City University, 2001.

[5] Smyth, B. and Cotter, P. A Personalized Television Listings Service. *Communications of the ACM*, 43(8), 2000, 107-111.

XP-002289742

P.D. 03-05-2001
P. 1-4 (4)

# FÍSCHLÁR: AN ON-LINE SYSTEM FOR INDEXING AND BROWSING BROADCAST TELEVISION CONTENT

*N. E. O'Connor, S. Marlow, N. Murphy, A. F. Smeaton, P. Browne, S. Deasy, H. Lee and K. McDonald*

Centre for Digital Video Processing
Dublin City University, Glasnevin, Dublin 9, Ireland.
oconnorn@eeng.dcu.ie
http://lorca.compapp.dcu.ie/Video

## ABSTRACT

This paper describes a demonstration system which automatically indexes broadcast television content for subsequent non-linear browsing. User-specified television programmes are captured in MPEG-1 format and analysed using a number of video indexing tools such as shot boundary detection, keyframe extraction, shot clustering and news story segmentation. A number of different interfaces have been developed which allow a user to browse the visual index created by these analysis tools. These interfaces are designed to facilitate users locating video content of particular interest. Once such content is located, the MPEG-1 bitstream can be streamed to the user in real-time. This paper describes both the high-level functionality of the system and the low-level indexing tools employed, as well as giving an overview of the different browsing mechanisms employed.

## 1. INTRODUCTION

Applications and services based on digital video content are becoming more widespread. This trend is likely to continue as evidenced by the increasing use of intranet video streaming in the workplace, the introduction and subsequent take-up of DVD and digital TV, as well as the deployment of broadband telecommunications networks to the home. With the increasing amount of video information available, there exists a need for efficient management of this information on behalf of the provider and a complementary need for efficient access and navigation of the content by the end user.

The Centre for Digital Video Processing at Dublin City University is pursuing an on-going research effort to develop essential technologies required for efficient management of video content. The project concentrates on fully automatic video indexing processes addressing both shot-level and scene-level video segmentation. The Centre also addresses the provision of good video content navigation and browsing support for end-users, which we believe to be an equally important aspect of video management. The work of the Centre to date is demonstrated via the web-based Físchlár[1] system.

In this paper we describe the high-level system functionality of Físchlár, the low-level indexing processes and the various browsing/navigation interfaces we have developed which support the provision of this functionality. An overview of the entire Físchlár system is presented in Section 2 which also describes the user mechanisms for recording (i.e. video capture) and browsing. Section 3 describes the various visual indexing tools we have implemented in the system. The six different browsing interfaces we have developed are outlined in Section 4. Finally, our plans for future work with the system are presented in Section 5.

## 2. SYSTEM OVERVIEW

Físchlár is a web-based demonstration system which allows users to (i) browse today's and tomorrow's television listings, (ii) select programmes to be recorded, analysed and indexed, (iii) view the visual index created by the system's indexing tools and (iv) select content, based on the index, and have it streamed to them in real-time [1]. The video server used in the system can store approximately 400 hours of video content, whilst the streaming technology employed supports 100 concurrent users.

Users can select programmes from eight terrestrial public broadcast channels. Television schedules can be viewed by channel, programme genre (e.g. comedy, drama, sports, etc.) or day (i.e. today or tomorrow). Most recently, a personalised listing service was introduced in order to offer programme recommendations based on user feedback on previously recorded content [2]. When a programme is recorded, it is captured in MPEG-1 format and stored on

The work described in this paper was funded by the National Software Directorate of Ireland with additional support from the Research Institute for Networks and Communications Engineering (RINCE).

[1] The name Físchlár is derived from two words in the Irish language: *fís* meaning dream or vision and *chlár* meaning programme

the system's video server. This MPEG-1 video bitstream is then analysed using a set of indexing tools in order to create a visual index for the content (see section 3).

Once the visual index has been created it can be presented to the user in the browse/playback section of Físchlár. In the browse/playback section, the list of recorded programmes currently stored by the system is displayed. The user can browse this list by date, channel, category (i.e. genre) or personalised recommendation. Once a programme is selected for viewing, its visual index is presented to the user for further browsing at the level of shots or scenes. The visual index for each programme consists of a set of shot boundaries and associated keyframes, possibly grouped by scene or subject. A number of different interfaces has been developed, which allow a user to browse this visual index in order to locate video segments of particular interest (see section 4). Once such a segment has been located, the MPEG-1 bitstream for that part of the programme can be streamed to the user. An example of the browse/playback functionality of Físchlár is illustrated in Figure 1.

**Fig. 1**. Browse and playback in Físchlár

## 3. INDEXING TOOLS

In this section, the different video indexing tools we have developed and integrated into Físchlár are described.

### 3.1. Shot-boundary detection and keyframe extraction

The core technology in any video indexing system is shot-boundary detection. We have investigated a number of different shot boundary detection algorithms [3, 4, 5]. The first algorithm investigated (and the algorithm currently employed in the "live" version of Físchlár) uses YUV colour histograms [3]. A histogram with 192 bins is computed for each image and compared with the previous image using the cosine distance similarity measure. A dynamic thresholding operation which adapts to the characteristics of the content being analysed is employed in order to detect shot boundaries. This approach works well for shot cuts but may lead to over segmentation in the case of fades or dissolves. For this reason, a shot boundary detection algorithm based on edge detection was investigated [4, 5]. A Sobel edge detector is applied to each decoded luminance image and the number of differing edge pixels between two successive images is calculated. Again, a thresholding process is employed in order to detect fades and dissolves. In an attempt to make the shot boundary detection algorithm as computationally efficient as possible, an approach based on counting MPEG-1 macro-block types was also investigated [5]. This approach detects when the number of Intra coded blocks rises above a pre-determined threshold signalling a shot boundary.

In order to aid our investigations, an evaluation baseline consisting of eight hours of manually indexed television content was employed. This base-line consists of different types of television content such as news programmes, soap operas, etc [3]. Every shot boundary detection algorithm we develop is applied to this base-line allowing their relative performance on a large test corpus to be evaluated. Using this baseline, work is already underway to investigate combining the three approaches outlined above into a unified approach [5].

Given shot boundaries for a programme, the next step is to extract a representative keyframe for each shot. The approach used selects a keyframe based on its similarity (using the cosine distance metric) to the average histogram calculated over the entire shot [3]. This approach was compared to approaches which simply select the first, middle or last video frame in a shot and was found to result in subjectively better representative keyframes, although this improvement is marginal.

### 3.2. Semantic boundary detection

Whilst extracting a key frame for each shot gives an overview of the contents of the video, typically this corresponds to a large amount of information which must be presented to the user. In general, people remember different events after viewing video content (and indeed think in terms of events during the information retrieval process) [6]. An event can be a dialog, action scene, news story or any other series of shots that are semantically related. For this reason we have developed a number of semantic boundary detection tools. A semantic boundary is defined as the boundary between two semantic units where a semantic unit is a series of consecutive shots that are related by some common theme or location [7].

In order to perform scene-level analysis of the content, a shot clustering algorithm has been developed. The algorithm we have implemented is based on the temporally constrained clustering approach of Rui *et al* [8]. The main

difference between our approach and that of Rui *et al* is the choice of features used for each shot. We use a single feature corresponding to the average histogram of the shot, rather than the multiple feature approach of Rui *et al*. We have found that this approach has worked well for our preliminary investigations but recognise that it will need to be extended in the future. The result of shot clustering is a set of groups consisting of visually similar shots. The relative temporal location of shots across groups is then analysed and temporal overlaps are detected in order to detect rudimentary scene boundaries [8].

The output groups of shots have also been used in a semantic boundary detection context in order to segment individual stories in Irish news programmes. The approach taken is to attempt to identify groups of shots corresponding to an anchor person. To this end, a number of heuristics based on the statistics of the groups are used. The statistics considered are the mean and standard deviation of the shot similarity measure, the mean and standard deviation of the temporal distance between shots, the number of shots and the mean shot length. Four rules are applied which successively eliminate groups as potential anchor person groups to finally settle on the set of groups which most probably contain an anchor person. This approach is designed to allow for news programmes with multiple news readers. The rules employed attempt to encapsulate the following characteristics of anchor person shots and groups: (i) anchor person groups tend to be larger than most other groups due to the fact that there are many similar shots contained within the entire news programme, (ii) anchor person shots tend to be longer than most other shots in a news programme, (iii) anchor person shots tend to have a global re-occurrence throughout a news programme whereas other shots are localised in time, (iv) anchor person shots tend to be extremely similar to each other. Some illustrative results of anchor person shot detection are illustrated in Figure 2.

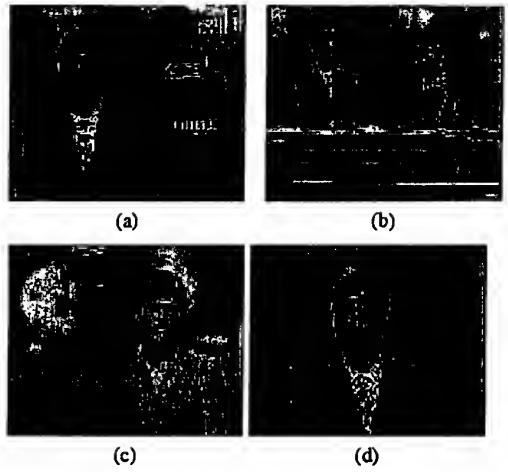



**Fig. 2.** Sample results of anchor person shot detection in a news programme

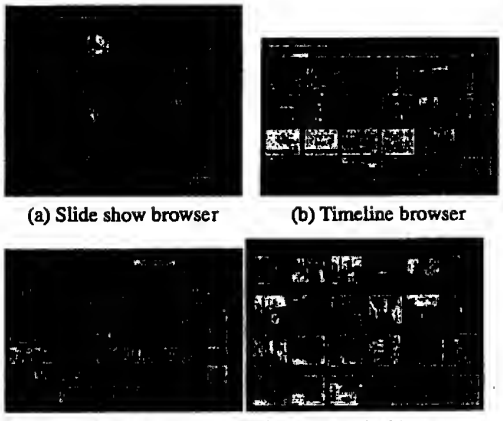
## 4. BROWSING INTERFACES

Various features of video querying and browsing interfaces are introduced and categorised in [9]. The particular design methodology we employed in developing a selection of keyframe-based video browsing interfaces for Físchlár is discussed in detail in [10]. As such, in this section we simply present a high-level overview of the interfaces we have developed.

In the scroll bar browser, the user simply scrolls up and down through all available keyframes which are arranged left to right, top to bottom in order of increasing temporal location in the programme. The advantage of this interface is that it is easy to use. However, such an approach can result in "information overload" for users due to the large number of keyframes associated with video content of any substantial length. In the slide show browser (see Figure 3(a)), keyframes are automatically displayed to the user one by one at rate of 2 per second (approx.). The user can also manually step forwards and backwards through the set of keyframes. A timeline indicator below the keyframes indicates the current temporal location in the programme. The main advantage of this interface is that it provides a summary of the content to the user. The main disadvantages are that typically this summary takes too long and that it is easy for a user to lose the context of what he/she is watching.

The timeline browser (see Figure 3(b)) presents a fixed number (24) of keyframes on one screen. The user can move between screens, and thus browse different sets of keyframes by selecting the associated temporal segment on the timeline bar. The timeline bar provides temporal orientation for users since it is segmented in proportion to the time spanned by a set of keyframes. A ToolTip indicating the exact start and end time of each segment is also provided. Feedback indicates that our users have found this interface attractive and easy to use. The initial screen of the overview/detail browser displays a small number of significant keyframes (see Figure 3(c)). A more detailed view of the video can be obtained on the second screen of this browser which presents the timeline browser to the user. The overview keyframes are selected based on the results of the scene-level analysis in the generic case, and on the results of anchor person detection in the specific case of news programmes. In the hierarchical browser, keyframes are grouped into a hierarchical tree structure which the user can navigate by moving up or down levels in the hierarchy

(see Figure 3(d)). The highest level consists of a small set of keyframes representative of the entire programme. The selection of these keyframes implicitly defines a temporal segmentation or grouping of the set of keyframes. Subsequent levels contain further segmentations of the previous level. This approach has previously been presented in [11]. Currently in Físchlár, the grouping which forms the temporal segmentation at each level is pre-defined and is not based on the results on semantic boundary detection.



(a) Slide show browser (b) Timeline browser

(c) Overview/detail browser (d) Hierarchical browser

**Fig. 3.** Browsing interfaces

## 5. CONCLUSIONS AND FUTURE WORK

The Físchlár system is currently used by a small set of technically oriented users. Preparations are underway to extend this user group to include both technical and non-technical users, corresponding to undergraduate and postgraduate students in the University. This would constitute a more representative user group and facilitate rigorous usability studies of our system.

To date, all indexing tools employed in the system work purely on the visual aspect of the video content. This is usually sufficient for tasks such as shot boundary detection and keyframe extraction. However, semantic boundary detection would benefit considerably from some analysis of the audio signal. For this reason, it is intended to develop a set of audio analysis tools which can be combined with our existing tools in order to perform scene-level and eventually event/object-level analysis with a view to aiding the detection of semantic boundaries. Tools such as silence detection, speech vs music classification and speaker segmentation are already being developed.

## 6. REFERENCES


[1] H. Lee et al, "The físchlár digital video recording, analysis, and browsing system," in *Proc. Content-based Multimedia Information Access (RIAO'2000)*, Paris, France, 12-14 Apr. 2000.

[2] B. Smith and P. Cotter, "A personalized television listings service," *Communications of the ACM*, vol. 43, no. 8, pp. 107–111, 2000.

[3] C. O'Toole et al, "Evaluation of automatic shot boundary detection on a large video test suite," in *Proc. The Challenge of Image Retrieval - 2nd UK Conference on Image Retrieval (CIR'99)*, Newcastle, UK, 25-26 Feb. 1999.

[4] A. Smeaton et al, "An evaluation of alternative techniques for automatic detection of shot boundaries in digital video," in *Proc. Irish Machine Vision and Image Processing Conference (IMVIP'99)*, Dublin, Ireland, 8-9 Sep. 1999.

[5] P. Browne et al, "Evaluating and combining digital video shot boundary detection algorithms," in *Proc. Irish Machine Vision and Image Processing Conference (IMVIP'2000)*, Belfast, Northern Ireland, 2000.

[6] A. Hanjalic J. Biemond, R. Lagendijk, "Automatically segmenting movies into logical story units," in *Proc. of the Third International Conference VISUAL '99*, Amsterdam, Netherlands, June 1999, pp. 229–236, Springer-Verlag.

[7] D. Petkovic P. Aigrain, H. Zhang, "Content-based representation and retrieval of visual media: A state-of-the-art review," *Multimedia tools and applications*, vol. 3, pp. 179–202, 1996.

[8] S. Mehrotra Y. Rui, T. S. Huang, "Constructing table-of-content for videos," *Multimedia Systems*, vol. 7, pp. 359–368, 1999.

[9] H. Lee et al, "User-interface issues for browsing digital video," in *Proc. 21st Annual Colloquium on IR Research (IRSG 99)*, Glasgow, UK, 19-20 Apr. 1999.

[10] H. Lee et al, "Implementation and analysis of several keyframe-based browsing interfaces to digital video," in *Proc. 4th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2000)*, Lisbon, Portugal, 18-20 Sep. 2000.

[11] H. Zhang et al, "Video parsing, retrieval and browsing: an integrated and content-based solution," in *Proc. of ACM International Conference on Multimedia '95*, New York, 1995, pp. 15–24.